Surrogate Modeling and Parameter Estimation Exercise

Simon Scheidegger

Department of Economics, University of Lausanne

Wednesday, March 25th, 2025

Overview of the Exercise

- We have a two-dimensional function $f(a, x) = (x - a)^3 + (x - a)$.
- Domain:
  - $a \in [1, 2]$
  - $x \in [0, 4]$
- We want to create a surrogate using Gaussian Process regression (GPyTorch).
- Then we mimic structural estimation by finding which parameter a corresponds to an empirical observation $y = 1.0$.

Task 1: Building the Surrogate

1. Define and sample the function $f(a,x)$:

$$f(a,x) = (x-a)^3 + (x-a).$$

2. Generate $N = 40$ random points in $[1,2] \times [0,4]$.
3. Evaluate f at these points to obtain training data.
4. Fit a Gaussian Process model to these data:
    - Use a constant mean function.
    - Use an RBF kernel.
    - Optimize the GP hyperparameters by maximizing log marginal likelihood.

Task 2: Parameter Estimation

- Hypothetical observation: $y = 1.0$.
- We want to discover which parameter $a \in [1, 2]$ makes the surrogate function's prediction close to $y = 1.0$.
- We define a criterion (e.g., sum of squared errors):

$$\text{Mismatch}(a) = \sum_{x \in [0,4]} \left(\hat{f}(a, x) - 1.0\right)^2.$$

- We search over a on a grid (or use an optimizer) to find the minimizing value.

Key Points and Results

- The GP surrogate approximates the true function well in the sampled domain.
- By minimizing the distance to $y = 1.0$, we identify which parameter a is consistent with that empirical observation (in a least-squares sense).
- This approach generalizes to higher-dimensional problems and more sophisticated models.

Conclusion

- Gaussian Processes provide a flexible surrogate for complicated functions.
- Structural estimation can be approximated by searching the parameter space within the trained surrogate.
- The final code demonstrates how to implement and solve this in Python.
- For further exploration, consider varying kernel types or including prior knowledge in the GP.

Thank you!